KONGERIKET NORGE
The Kingdom of Norway

PCT/NO 03/00088
10/507434
10 Rec'd PCT/PTO 14 SEP 2004

REC'D 08 APR 2003
WIPO PCT

Bekreftelse på patentsøknad nr
Certification of patent application no

2002 1298

Det bekreftes herved at vedheftede dokument er nøyaktig utskrift/kopi av ovennevnte søknad, som opprinnelig inngitt 2002.03.15

It is hereby certified that the annexed document is a true copy of the above-mentioned application, as originally filed on 2002.03.15

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

2003.03.21

Freddy Strømmen

Freddy Strømmen
Seksjonsleder

Line Reum

Line Reum

PATENTSTYRET®
Styret for det industrielle rettsvern

BL 130 - 5.000 - 01.2001 - Hegro Grafisk AS

BEST AVAILABLE COPY

1d

EKR 15.03.2002 E24418

Bjørn H. Lindqvist
Ankerveien 13B
0782 OSLO

Oppfinner: Søkeren

Methods for peptide and protein display in nucleic acid chip-arrays

METHODS FOR PEPTIDE AND PROTEIN DISPLAY IN NUCLEIC ACID CHIP-ARRAYS

FIELD OF THE INVENTION

The present invention relates to methods of creating a nucleic acid based protein display chip-array, where a diverse population of peptides or proteins is displayed.

BACKGROUND OF THE INVENTION

Recent advances in chip-array technology have led to high through-put analysis of gene expression at the level of transcription. To achieve functional array based genomics, however, the proteins encoded by the genes in the arrays should also be present and displayed and presented (Cahill, 2000, Irving and Hudson, 2000). The present invention provides all said features by presenting a method where the protein encoded by the gene is indeed present, displayed and presented. This has been made possible by taking advantage of a published observation (Chattaraj and Inman, 1974, Thomas, 1974) that certain virus are able to release their chromosome from their capsids while remaining covalently attached to their tail or head and combine this observation with phage display and nucleic acid chip-array technologies. The present invention provides thereby a new and more simple method in creating a nucleic acid based protein display chip-array.

Nucleic acid chip-array primarily DNA arrays, may contain thousands of individual nucleic acid addressable entries on a small surface (as small as one square cm) fabricated by robotic procedures involving fine mechanics and special chemistries. Phage display is a process by which a peptide or a protein is expressed – almost invariably as an exterior fusion to the surface protein of the phage particle in such a way that the peptide or protein sequence can be deduced from the coding DNA or RNA sequence residing in the phage particle or its transductant.

US Patent Application No: 6,207,446 describes methods for selection of protein molecules that make use of RNA-protein fusions, where a microchip comprises an array of immobilized single stranded nucleic acids, wherein said nucleic acid is being hybridised to RNA-protein fusions.

US Patent Application No: 6,194,550 describes a method called SPERT (Systemic Polypeptide Evolution by Reverse Translation), where a single stranded nucleic acid is immobilised on a solid support. A ribosomal complex is added and the protein corresponding to the immobilised gene sequence is then synthesized. The ribosome complex is removed and the solid support contains an immobilized nucleic acid–protein complex.

The present invention differs from prior approaches in that the proteins/peptides encoded by the genes/cDNA in the chip-arrays are present and presented without the extra steps of transcription and translation of the cDNA on the chip-array, since the present invention creates a protein display array as if the proteins had been formed by transcription and translation of the cDNA in the original library array. But instead the transcription and translation of the cDNA occurred during the propagation of the phage library.

Current protein array assemblies involve automation by mechanical and chemical deposition of proteins in array formats in (de Wildt et al. 2000, MacBeath and Schreiber 2000). In these cases the proteins may become abused when bound to the solid array surface. The hybridization step according to the present invention avoids this problem by leaving the protein/peptid displaying virion scaffold in solution.

SUMMARY OF THE INVENTION

The present invention relates to methods for producing coded nucleic acid chip-arrays which display a diverse population of peptides or protein. The nucleic acid coded protein or peptide display chip-arrays described in the present invention can be used in functional genomics, proteomics and in protein or peptide identification of relevance for

the exploration of therapeutic drugs as well as for search of new diagnostic procedures.

The nucleic acid coded protein or peptide display chip-array technology described here has the important features of self-assembly and auto-addressing a protein product to its gene while keeping the displayed protein/peptide in solution. The chip-array can be reused after removal of the hybridised DNA-virion-fusion complexes by nucleic acid denaturation.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 A provides an electronic microscope picture of a phage P4 naked DNA-virion protein complex.

Fig. 1B depicts electron microscope picture of lambda tail : : DNA complexes.

Fig. 1C provides a drawing of ss-DNA tail complex displaying a protein.

Fig. 2 depicts a drawing of a magnification of one display virion unit in the process of being deposited by hybridisation in a DNA-array after DNA:DNA hybridisation in the array. The bottom drawing shows a virion-DNA display complex hybridising to a DNA-Array entry.

DEFINITIONS

As used herein, the term "virion" refers to the bacteriophage particle

As used herein, the term "naked DNA" refers to the free bacteriophage DNA released from the virion.

As used herein, the term "naked DNA tail/head complexes" refers to the released and naked bacteriophage DNA attached to the virion proteins making up the "tail" and the "head".

As used herein, the term "naked nucleic acid virion protein display" denotes the naked DNA tail/head complexes displaying a peptide or protein.

As used herein the term "viron::protein" is defined as virion displaying a peptide or protein.

As used herein the term DNA virion fusion complexes refers to the naked

DNA tail/head complexes displaying a peptide or protein

As used herein the term "cis-capture" is defined as the linking of peptide/protein or a protein complex to the very same template that encoded these protein entities.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to methods for producing coded nucleic acid chip-arrays which display a diverse population of peptides or protein. The nucleic acid coded protein or peptide display chip-arrays described in the present invention can be used in functional genomics, proteomics and in protein or peptide identification of relevance for the exploration of therapeutic drugs as well as for search of new diagnostic procedures.

The present invention describes methods for protein/peptide display in nucleic acid chip-arrays by integrating phage display in chip array formats. The method takes advantage of a published observation (Chattaraj and Inman, 1974) that, under certain conditions, more than 95 % of the chromosomes of certain phages (e.g. lambda, P2 and P4) can be released naked from the phage capsids while remaining covalently attached - by one of their ends (always the same end) - to the top of their respective tail sometimes associated with the head (here referred to as a naked nucleic acid-virion protein complex and exemplified by phage P4 (Lindqvist unpublished observation), fig. 1). This observation makes it possible to generate naked nucleic acid-virion protein display complexes in which the individual nucleic acid template is freely and covalently linked to the very same virion proteins it coded for. This *cis*-capture of the virion proteins by its naked nucleic acid template makes it possible to use the converted phage display libraries in combinatorial display chip-array formats. This can be achieved by allowing the naked nucleic acid-virion protein display complexes to be deposited by nucleic acid hybridization to their corresponding mRNA or cDNA which separately have been prepared and used for the fabrication of a nucleic acid chip-array. The hybridization step will therefore function as a 'search engine' and a 'delivery robot' by automatically positioning the protein to its own gene (sometimes related gene due to cross hybridization) in the

chip-array by gently depositing the virion protein display complex to the solid surface.

This procedure thus creates a protein display chip-array as if the proteins had been formed by transcription and translation of the cDNA in the chip-array. But instead the transcription and translation of the cDNA occured during the propagation of the phage library in *E. coli.*

Chip arrays may comprise of two or three dimensional arrays including particle based arrays kept in solution.

Creating a nucleic acid based protein display chip-array

In a preferred embodiment lambda is applied since lambda can display functional proteins fused both to its tail or head, by utilizing the V or the D proteins respectively (Maruyama et al. 1994, Dunn, 1995, 1996, Sternberg and Hoess 1995, Mikawa et al 1996). V consists of 246 amino acids and it is present as 192 copies in the form of 32 hexameric discs. Since the most efficient chromosome *cis*-capture occurs with the tail under the present protocols, lambda particles displaying a library of peptides or cDNA encoded proteins fused to the C-terminal part of the major tail protein V are used for the conversion into naked nucleic acid-virion protein display complexes. A protein display chip-array can then be created, where samples of individual phage stocks are used or individual plaques originating from a phage display library are picked and and then treated to yield pure naked DNA-tail/head display complexes followed by their deposition in an chip-array format by specifically cross-linking the DNA part to a solid support. The tail will then function as a giant free moving scaffold in which peptides or proteins fused to the V protomers are displayed. The steps of array assembly should be amendable to automation in a similar way described by de Wildt et al. (2000) or MacBeath and Schreiber (2000). The nature of any protein in the array interacting with a defined target or displaying a measurable function can easily be deduced by sequencing its stored phage replica, a procedure which would not differ from that used in standard phage display.

Another preferred embodiment is to hybridise the naked DNA-tail/head display complexes against a complementary cDNA library chip-array, in such a way that each member of the peptide/protein library will hybridise to its homologue cDNA sequence

if present in the array (Fig 2). In order to avoid cross-hybridisation the cDNA present in the array should be a part of a vector other than lambda. Before hybridisation, however, the DNA bound to the tail/head has to be made single stranded by exonucleases such as either exoIII (3') or lambda exo (5') degradation, one of the ends of the chromosome is hidden in the tail which protects it from exo-nuclease attack. This treatment generates ssDNA-tail/head display complexes as well as removes free DNA which failed to *cis*-capture the tail/head display complex. Since the phage library is amplified after its construction, a variable number of copies (the size of which will depend on the size of the library and the relative frequency of the individual mRNA/cDNA molecules converted to display format) of each unique DNA-tail/head display complex should be present and accessible for hybridisation. Again, the DNA-DNA hybridisation step will therefore function as a 'search engine' and a 'delivery robot' by automatically positioning the protein to its own (or related) gene in the array by gently depositing the tail/head display complex to the solid surface. This procedure thus creates a protein display array as if the proteins had been formed by transcription and translation of the cDNA in the original library array. But instead the transcription and translation (assuming an equal and constant translation of the cDNA during the phage propagation, the protein deposited in each position of the array should reflect the relative frequency of mRNA molecules in the cell at the time of harvest (and not necessarily the protein amount) of the cDNA occured during the propagation of the phage library.

The efficiency of display in the chip-array will be a function of the ratio between the size of the library and the number of cDNA clones in the array. If an array consists of 1000 cDNA clones originating from an organism of approximately 5000 genes the DNA-tail/head display library must carry a cDNA representative of each of those genes in order to be fully "expressed". The lambda V display vectors should easily handle up to 10^7 independent clones and a $5x10^3$ cDNA library can be made to consist of up to 10^8 copies or more of each cDNA clone after amplification. Therefore, each cDNA spot in the array can potentially contain up to 10^8 or more copies of its encoded protein (assuming that one to two protein molecule are expressed as a fusion of V per particle). This is known to be the case for functional proteins expressed as fusions of V whereas peptides can be expressed as fusions of each V protomer (Maruyama et al 1994, Dunn,

1995, 1996) thereby increasing the presentation by a factor by 192 in each spot. In either case the amount of displayed peptide/protein in each spot will fall in the pg and sub-pmole range.

Use of the method

The nucleic acid coded protein or peptide display chip-arrays described here can be used in functional genomics, proteomics and in protein or peptide identification of relevance for the exploration of therapeutic drugs as well as in the search for new diagnostic procedures. The chip-array can be reused after removal of the hybridised DNA-virion-fusion complexes by nucleic acid denaturation. The naked nucleic acid-virion protein display complex can also be used as a scaffold for bi-functional display after nucleic acid hybridisation of two different display complexes.

REFERENCES

1.Cahill, D. Protein arrays: a high-throughput solution for proteonomics research. Proteonomics: A trends Guide. (2000) July, 47-51

2.Chattaraj and Inman. Location of DNA ends in P2, 186, P4 and lambda bacteriophage heads. J. Mol. Biol. (1974) 87, 11-22

3. Davis, R. W., Simon, M. and N. Davidson. In Methods in Enzymology (1971) 21, part D 414-428

4. DeWildt, R. M., Mundy, C. R., Gorick, B. D. and I. M Tomlinson. Antibody arrays for high-throuput screening of antibody-antigen interactions. Nat. Biotechnology (2000) 18 989-994

5. Dunn I. S. Assembly of functional bacteriophage lambda virions incorporating C-terminal peptide or protein fusion with the major tial protein. J. Mol. Biol (1995) 248, 497-506

6. Dunn I. S. Total modification of the bacteriophage lambda tail tube major subunit protein with foreign peptides. Gene (1996) 183, 15-21

7. Irving, R. A. and P. Hudson. Proteins emerge from disaaray. Nat. Biotchnology (2000) 18, 932-933

8. MacBeath, G and S. L. Schreiber. Printing Proteins as Microarrays for High-throughput Function Determination, Science (2000) 289, 1760-1763

9. Maruyama IN, Maruyama HI, S. Brenner. Lambda foo: a lambda phage vector for the expression of foreign proteins. Proc Natl Acad Sci U S A. (1994) Aug 16;91(17):8273-7.

10. Mikawa Y G., Maruyama, I. N. and S. Brenner. Surface display of proteins on

bacteriophage lambda heads. J. Mol Biol (1996) 262, 21-30.

11. Sternberg, N and R. H. Hoess. Display of peptides and proteins on the surface of bacteriophage lambda. Proc. Nat. Acad. Sci (USA) (1995) 28, 1609-1613

12. Thomas, J.O. Chemical linkage of the tail to the right-hand end of bacteriophage lambda DNA. J. Mol. Biol. (1974) 87, 1-9

CLAIMS

1.

A method of preparing covalently linked naked nucleic acid-protein display complexes from virus particles, wherein the nucleic acid part of the complex is linked to a support or where optionally said display complex is kept in a solution

2.

A method of claim 1 comprising at least the steps of:

a) treating a fresh prepared virus preparation with cross linking chemical agents producing covalent linked naked nucleic acid-virion protein display complexes,

b) coupling of the naked nucleic acid-virion protein display complexes by chemical cross linking agents to a solid support, or by hybridising of the naked nucleic acid-virion protein display complexes and thereby,

c) positioning the displayed protein/peptide to its own gene or related gene(s).

3.

A method of claim 1 and 2 wherein the said naked nucleic acid-virion protein display complexes are fabricated from head and tail containing virus particles by covalent cis-capture of the virion by the naked nucleic acid chromosome.

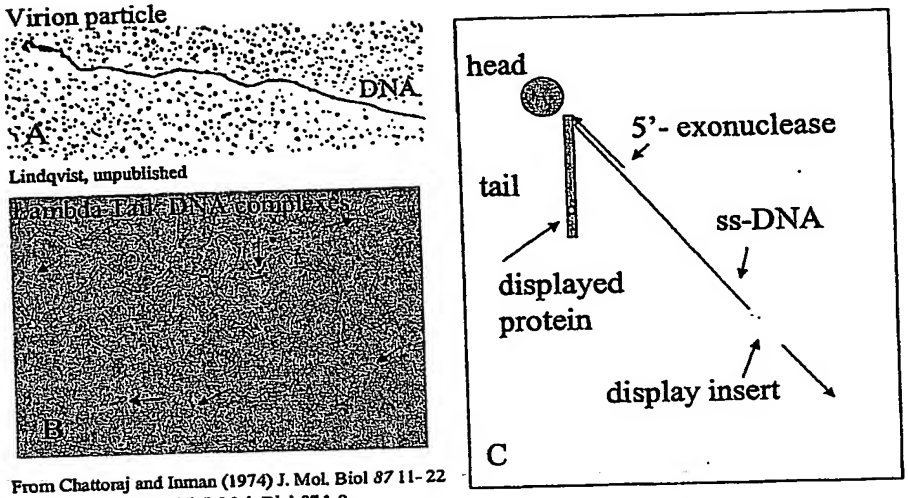
4.

A method of claims 1-3 wherein a purified stock/plaque of virus particles displaying the protein/peptide is the source of said naked nucleic acid-virion protein display complexes.

5.

A method of claims 1-3 wherein viral display libraries of proteins/peptides are the source of said naked nucleic acid-virion protein display complexes.

Figure 1.



From Chattoraj and Inman (1974) J. Mol. Biol *87* 11-22
see also Thomas (1974) J. Mol. Biol *87* 1-9

Figure 2.

Integrating Phage Display and Nucleic Acid Array Technologies

- Creating a Naked Virion-DNA Display Complex
- A DNA array
- A Virion-DNA Display Complex Hybridizing to a DNA-Array Entry